Household Data ("A" tables, monthly; "D" tables, quarterly)

COLLECTION AND COVERAGE

Statistics on the employment status of the population and related data are compiled by BLS using data from the Current Population Survey (CPS). This monthly survey of households is conducted for BLS by the U.S. Census Bureau through a scientifically selected sample designed to represent the civilian noninstitutional population. Respondents are interviewed to obtain information about the employment status of each member of the household 16 years of age and older. The inquiry relates to activity or status during the calendar week, Sunday through Saturday, that includes the 12th day of the month. This is known as the "reference week." Actual field interviewing is conducted in the following week, referred to as the "survey week."

Each month, about 60,000 occupied units are eligible for interview. Some 4,500 of these households are contacted but interviews are not obtained because the occupants are not at home after repeated calls or are unavailable for other reasons. This represents a noninterview rate for the survey that ranges between 7 and 8 percent. In addition to the 60,000 occupied units, there are about 12,000 sample units in an average month that are visited but found to be vacant or otherwise not eligible for enumeration. Part of the sample is changed each month. The rotation plan, as will be explained later, provides for three-fourths of the sample to be common from one month to the next, and one-half to be common with the same month a year earlier.

CONCEPTS AND DEFINITIONS

The concepts and definitions underlying labor force data have been modified, but not substantially altered, since the inception of the survey in 1940; those in use as of January 1994 are as follows:

Civilian noninstitutional population. Included are persons 16 years of age and older residing in the 50 States and the District of Columbia who are not inmates of institutions (for example, penal and mental facilities, homes for the aged), and who are not on active duty in the Armed Forces.

Employed persons. All persons who, during the reference week, (a) did any work at all (at least 1 hour) as paid employees, worked in their own business, profession, or on their own farm, or worked 15 hours or more as unpaid workers in an enterprise operated by a member of the family, and (b) all those who were not working but who had jobs or businesses from which they were temporarily absent because of vacation, illness, bad weather, childcare problems, maternity or paternity leave, labor-management dispute, job training, or other family or personal reasons, whether or not they were paid for the time off or were seeking other jobs.

Each employed person is counted only once, even if he or she holds more than one job. For purposes of occupation and industry classification, multiple jobholders are counted in the job at which they worked the greatest number of hours during the reference week.

Included in the total are employed citizens of foreign countries who are temporarily in the United States but not living on the premises of an embassy. Excluded are persons whose only activity consisted of work around their own house (painting, repairing, or own home housework) or volunteer work for religious, charitable, and other organizations.

Unemployed persons. All persons who had no employment during the reference week, were available for work, except for temporary illness, and had made specific efforts to find employment sometime during the 4-week period ending with the reference week. Persons who were waiting to be recalled to a job from which they had been laid off need not have been looking for work to be classified as unemployed.

Duration of unemployment. This represents the length of time (through the current reference week) that persons classified as unemployed had been looking for work. For persons on layoff, duration of unemployment represents the number of full weeks they had been on layoff. Mean duration is the arithmetic average computed from single weeks of unemployment; median duration is the midpoint of a distribution of weeks of unemployment.

Reason for unemployment. Unemployment also is categorized according to the status of individuals at the time they began to look for work. The reasons for unemployment are divided into five major groups: (1) **Job losers**, comprising (a) persons *on temporary layoff*, who have been given a date to return to work or who expect to return within 6 months (persons on layoff need not be looking for work to qualify as unemployed), and (b) **permanent job losers**, whose employment ended involuntarily and who began looking for work; (2) **Job leavers**, persons who quit or otherwise terminated their employment voluntarily and immediately began looking for work; (3) Persons who completed temporary jobs, who began looking for work after the jobs ended; (4) Reentrants, persons who previously worked but who were out of the labor force prior to beginning their job search; and (5) *New entrants*, persons who had never worked. Each of these five categories of the unemployed can be expressed as a proportion of the entire civilian labor force; the sum of the four rates thus equals the unemployment rate for all civilian workers. (For statistical presentation purposes, "job losers" and "persons who completed temporary jobs" are combined into a single category until seasonal adjustments can be developed for the separate categories.)

Jobseekers. All unemployed persons who made specific efforts to find a job sometime during the 4-week period preceding the survey week are classified as jobseekers. Jobseekers do not include persons classified as on temporary layoff, who, although often looking for work, are not required to do so to be classified as unemployed. Jobseekers are grouped by the methods used to seek work. Only active methods—which have the potential to result in a job offer without further action on the part of the jobseeker—qualify as job search. Examples include going to an employer directly or to a public or private employment agency, seeking assistance from friends or relatives, placing or answering ads, or using some other active method. Examples of the "other" category include being on a union or professional register, obtaining assistance from a community organization, or waiting at a designated labor pickup point. Passive methods, which do not qualify as job search, include reading (as opposed to answering or placing) "help wanted" ads and taking a job training course.

Labor force. This group comprises all persons classified as employed or unemployed in accordance with the criteria described above.

Unemployment rate. The unemployment rate represents the number unemployed as a percent of the labor force.

Participation rate. This represents the proportion of the population that is in the labor force.

Employment-population ratio. This represents the proportion of the population that is employed.

Not in the labor force. Included in this group are all persons in the civilian noninstitutional population who are neither employed nor unemployed. Information is collected on their desire for and availability to take a job at the time of the CPS interview, job search activity in the prior year, and reason for not looking in the 4-week period prior to the survey week. This group includes discouraged workers, defined as persons not in the labor force who want and are available for a job and who have looked for work sometime in the past 12 months (or since the end of their last job if they held one within the past 12 months), but who are not currently looking because they believe there are no jobs available or there are none for which they would qualify.

Persons classified as not in the labor force who are in the sample for either their fourth or eighth month are asked additional questions relating to job history and workseeking intentions. These latter data are available on a quarterly basis.

Occupation, industry, and class of worker. This information for the employed applies to the job held in the reference week. Persons with two or more jobs are classified in the job at which they worked the greatest number of hours. The unemployed are classified according to their last job. Beginning in 2003, the occupational and industrial classification of CPS data is based on the 2002 Census

Bureau occupational and industrial classification systems which are derived from the 2000 Standard Occupational Classification (SOC) and the 2002 North American Industry Classification System (NAICS). (See the following section on historical comparability for a discussion of previous classification systems used in the CPS.)

The class-of-worker breakdown assigns workers to the following categories: Private and government wage and salary workers, self-employed workers, and unpaid family workers. Wage and salary workers receive wages, salary, commissions, tips, or pay in kind from a private employer or from a government unit. Self-employed persons are those who work for profit or fees in their own business, profession, trade, or farm. Only the unincorporated self-employed are included in the self-employed category in the classof-worker typology. Self-employed persons who respond that their businesses are incorporated are included among wage and salary workers because, technically, they are paid employees of a corporation. Unpaid family workers are persons working without pay for 15 hours a week or more on a farm or in a business operated by a member of the household to whom they are related by birth or marriage.

Multiple jobholders. These are employed persons who, during the reference week, either had two or more jobs as a wage and salary worker, were self-employed and also held a wage and salary job, or worked as an unpaid family worker and also held a wage and salary job. Excluded are self-employed persons with multiple businesses and persons with multiple jobs as unpaid family workers.

Hours of work. These statistics relate to the actual number of hours worked during the reference week. For example, persons who normally work 40 hours a week but were off on the Columbus Day holiday would be reported as working 32 hours, even though they were paid for the holiday. For persons working in more than one job, the published figures relate to the number of hours worked in all jobs during the week; all the hours are credited to the major job. Unpublished data are available for the hours worked in each job and for usual hours.

At work part time for economic reasons. Sometimes referred to as involuntary part time, this category refers to individuals who gave an economic reason for working 1 to 34 hours during the reference week. Economic reasons include slack work or unfavorable business conditions, inability to find full-time work, and seasonal declines in demand. Those who usually work part time must also indicate that they want and are available for full-time work to be classified as on part time for economic reasons.

At work part time for noneconomic reasons. This group includes those persons who usually work part time and were at work 1 to 34 hours during the reference week for a noneconomic reason. Noneconomic reasons include, for example: Illness or other medical limitations, childcare problems or

other family or personal obligations, school or training, retirement or Social Security limits on earnings, and being in a job where full-time work is less than 35 hours. The group also includes those who gave an economic reason for usually working 1 to 34 hours but said they do not want to work full time or are unavailable for such work.

Usual full- or part-time status. Data on persons "at work" exclude persons who were temporarily absent from a job and therefore classified in the zero-hours-worked category, "with a job but not at work." These are persons who were absent from their jobs for the entire week for such reasons as bad weather, vacation, illness, or involvement in a labor dispute. In order to differentiate a person's normal schedule from his or her activity during the reference week, persons also are classified according to their usual full- or part-time status. In this context, full-time workers are those who usually worked 35 hours or more (at all jobs combined). This group will include some individuals who worked less than 35 hours in the reference week for either economic or noneconomic reasons and those who are temporarily absent from work. Similarly, *part-time workers* are those who usually work less than 35 hours per week (at all jobs), regardless of the number of hours worked in the reference week. This may include some individuals who actually worked more than 34 hours in the reference week, as well as those who are temporarily absent from work. The *full-time labor force* includes all employed persons who usually work full time and unemployed persons who are either looking for full-time work or are on layoff from full-time jobs. The part-time labor force consists of employed persons who usually work part time and unemployed persons who are seeking or are on layoff from parttime jobs. Unemployment rates for full- and part-time workers are calculated using the concepts of the full- and parttime labor force.

White, black or African American, and Asian. These are terms used to describe the race of persons. Persons in these categories are those who selected that race group only. Persons in the remaining race categories—American Indian or Alaska Native, Native Hawaiian or Other Pacific Islanders, and persons who selected more than one race category—are included in the estimates of total employment and unemployment but are not shown separately because the number of survey respondents is too small to develop estimates of sufficient quality for monthly publication. In the enumeration process, race is determined by the household respondent. (See the following section on historical comparability for a discussion of changes beginning in 2003 that affected how people are classified by race.)

Hispanic or Latino ethnicity. This refers to persons who identified themselves in the enumeration process as being Spanish, Hispanic, or Latino. Persons whose ethnicity is identified as Hispanic or Latino may be of any race. (See the following section on historical comparability for a

discussion of changes beginning in 2003 that affected how people are classified by Hispanic or Latino ethnicity.)

Usual weekly earnings. Data represent earnings before taxes and other deductions, and include any overtime pay, commissions, or tips usually received (at the main job, in the case of multiple jobholders). Earnings reported on a basis other than weekly (for example, annual, monthly, hourly) are converted to weekly. The term "usual" is as perceived by the respondent. If the respondent asks for a definition of usual, interviewers are instructed to define the term as more than half the weeks worked during the past 4 or 5 months. Data refer to wage and salary workers (excluding all self-employed persons regardless of whether their businesses were incorporated) who usually work full time on their sole or primary job.

Median earnings. These figures indicate the value that divides the earnings distribution into two equal parts, one part having values above the median and the other having values below the median. The medians shown in this publication are calculated by linear interpolation of the \$50 centered interval within which each median falls. Data expressed in constant dollars are deflated by the Consumer Price Index for All Urban Consumers (CPI-U).

Never married; married, spouse present; and other marital status. These are the terms used to define the marital status of individuals at the time of interview. Married, spouse present, applies to husband and wife if both were living in the same household, even though one may be temporarily absent on business, on vacation, on a visit, in a hospital, etc. Other marital status applies to persons who are married, spouse absent; widowed; or divorced. Married, spouse absent relates to persons who are separated due to marital problems, as well as to husbands and wives who are living apart because one or the other was employed elsewhere or was on duty with the Armed Forces, or for any other reasons.

Household. A household consists of all persons—related family members and all unrelated persons—who occupy a housing unit and have no other usual address. A house, an apartment, a group of rooms, or a single room is regarded as a housing unit when occupied or intended for occupancy as separate living quarters. A **householder** is the person (or one of the persons) in whose name the housing unit is owned or rented. The term is never applied to either husbands or wives in married-couple families but relates only to persons in families maintained by either men or women without a spouse.

Family. A family is defined as a group of two or more persons residing together who are related by birth, marriage, or adoption; all such persons are considered as members of one family. Families are classified either as married-couple families or as families maintained by women or men without spouses. A family maintained by a woman or a man is one in which the householder is either single, widowed, divorced, or married, spouse absent.

HISTORICAL COMPARABILITY

Changes in concepts and methods

While current survey concepts and methods are very similar to those introduced at the inception of the survey in 1940, a number of changes have been made over the years to improve the accuracy and usefulness of the data. Some of the most important changes include:

- In 1945, the questionnaire was radically changed with the introduction of four basic employment questions. Prior to that time, the survey did not contain specific question wording, but, rather, relied on a complicated scheme of activity prioritization.
- In 1953, the current 4-8-4 rotation system was adopted, whereby households are interviewed for 4 consecutive months, leave the sample for 8 months, and then return to the sample for the same 4 months of the following year. Before this system was introduced, households were interviewed for 6 consecutive months and then replaced. The new system provided some year-to-year overlap in the sample, thereby improving measurement over time.
- In 1955, the survey reference week was changed to the calendar week including the 12th day of the month, for greater consistency with the reference period used for other labor-related statistics. Previously, the calendar week containing the 8th day of the month had been used as the reference week.
- In 1957, the employment definition was modified slightly as a result of a comprehensive interagency review of labor force concepts and methods. Two relatively small groups of persons classified as employed, under "with a job but not at work," were assigned to different classifications. Persons on layoff with definite instructions to return to work within 30 days of the layoff date, and persons volunteering that they were waiting to start a new wage and salary job within 30 days of interview, were, for the most part, reassigned to the unemployed classification. The only exception was the small subgroup in school during the reference week but waiting to start new jobs, which was transferred to not in the labor force.
- In 1967, more substantive changes were made as a result of the recommendations of the President's Committee to Appraise Employment and Unemployment Statistics (the Gordon Committee). The principal improvements were as follows:
- a) A 4-week job search period and specific questions on jobseeking activity were introduced. Previously, the questionnaire was ambiguous as to the period for jobseeking, and there were no specific questions concerning job search methods.
- b) An availability test was introduced whereby a person must be currently available for work in order to be classified as unemployed. Previously, there was no such requirement. This revision to the concept mainly affected students, who, for example, may begin to look for summer jobs in the spring

- although they will not be available until June or July. Such persons, until 1967, had been classified as unemployed but since have been assigned to the "not in the labor force" category.
- c) Persons "with a job but not at work" because of strikes, bad weather, etc., who volunteered that they were looking for work were shifted from unemployed status to employed.
- d) The lower age limit for official statistics on employment, unemployment, and other labor force concepts was raised from 14 to 16 years. Historical data for most major series have been revised to provide consistent information based on the new minimum age limit.
- e) New questions were added to obtain additional information on persons not in the labor force, including those referred to as "discouraged workers," defined as persons who indicate that they want a job but are not currently looking because they believe there are no jobs available or none for which they would qualify.
- f) New "probing" questions were added to the questionnaire in order to increase the reliability of information on hours of work, duration of unemployment, and self-employment.
- In 1994, major changes to the Current Population Survey (CPS) were introduced, which included a complete redesign of the questionnaire and the use of computer-assisted interviewing for the entire survey. In addition, there were revisions to some of the labor force concepts and definitions, including the implementation of some changes recommended in 1979 by the National Commission on Employment and Unemployment Statistics (NCEUS, also known as the Levitan Commission). Some of the major changes to the survey were:
- a) The introduction of a redesigned and automated questionnaire. The CPS questionnaire was totally redesigned in order to obtain more accurate, comprehensive, and relevant information, and to take advantage of state-of-the-art computer interviewing techniques.
- b) The addition of two, more objective, criteria to the definition of discouraged workers. Prior to 1994, to be classified as a discouraged worker, a person must have wanted a job and been reported as not currently looking because of a belief that no jobs were available or that there were none for which he or she would qualify. Beginning in 1994, persons classified as discouraged must also have looked for a job within the past year (or since their last job, if they worked during the year), and must have been available for work during the reference week (a direct question on availability was added in 1994; prior to 1994, availability had been inferred from responses to other questions). These changes were made because the NCEUS and others felt that the previous definition of discouraged workers was too subjective, relying mainly on an individual's stated desire for a job and not on prior testing of the labor market.

- c) Similarly, the identification of persons employed part time for economic reasons (working less than 35 hours in the reference week because of poor business conditions or because of an inability to find full-time work) was tightened by adding two new criteria for persons who usually work part time: They must want and be available for full-time work. Previously, such information was inferred. (Persons who usually work full time but worked part time for an economic reason during the reference week are assumed to meet these criteria.)
- d) Specific questions were added about the expectation of recall for persons who indicate that they are on layoff. To be classified as "on temporary layoff," persons must expect to be recalled to their jobs. Previously, the questionnaire did not include explicit questions about the expectation of recall.
- e) Persons volunteering that they were waiting to start a new job within 30 days must have looked for work in the 4 weeks prior to the survey in order to be classified as unemployed. Previously, such persons did not have to meet the job search requirement in order to be included among the unemployed.

For additional information on changes in CPS concepts and methods, see "The Current Population Survey: Design and Methodology," Technical Paper 63RV (Washington, U.S. Census Bureau and Bureau of Labor Statistics, March 2002), available on the Internet at <code>www.bls.census.gov/cps/tp/tp63.htm</code>; "Overhauling the Current Population Survey—Why is it Necessary to Change?," "Redesigning the Questionnaire," and "Evaluating Changes in the Estimates," <code>Monthly Labor Review</code>, September 1993; and "Revisions in the Current Population Survey Effective January 1994," in the February 1994 issue of this publication.

Noncomparability of labor force levels

In addition to the refinements in concepts, definitions, and methods made over the years, other changes also have affected the comparability of the labor force data.

- Beginning in 1953, as a result of introducing data from the 1950 census into the estimating procedures, population levels were raised by about 600,000; labor force, total employment, and agricultural employment were increased by about 350,000, primarily affecting the figures for totals and for men; other categories were relatively unaffected.
- Beginning in 1960, the inclusion of Alaska and Hawaii resulted in increases of about 500,000 in the population and about 300,000 in the labor force. Four-fifths of the labor force increase was in nonagricultural employment; other labor force categories were not appreciably affected.
- Beginning in 1962, the introduction of data from the 1960 census reduced the population by about 50,000 and labor force and employment by about 200,000; unemployment totals were virtually unchanged.

- Beginning in 1972, information from the 1970 census was introduced into the estimation procedures, increasing the population by about 800,000; labor force and employment totals were raised by a little more than 300,000; unemployment levels and rates were essentially unchanged.
- In March 1973, a subsequent population adjustment based on the 1970 census was introduced. This adjustment, which affected the white and black-and-other groups but had little effect on totals, resulted in the reduction of nearly 300,000 in the white population and an increase of the same magnitude in the black-and-other population. Civilian labor force and total employment figures were affected to a lesser degree; the white labor force was reduced by 150,000, and the black-and-other labor force rose by about 210,000. Unemployment levels and rates were not significantly affected.
- Beginning in January 1974, the method used to prepare independent estimates of the civilian noninstitutional population was modified to an "inflation-deflation" approach. This change in the derivation of the estimates had its greatest impact on estimates of 20- to 24-year-old men—particularly those in the black-and-other population—but had little effect on estimates of the total population 16 years and over. Additional information on the adjustment procedure appears in "CPS Population Controls Derived from Inflation-Deflation Method of Estimation," in the February 1974 issue of this publication.
- Effective in July 1975, as a result of the large inflow of Vietnamese refugees to the United States, the total and black-and-other independent population controls for persons 16 years and over were adjusted upward by 76,000—30,000 men and 46,000 women. The addition of the refugees increased the black-and-other population by less than 1 percent in any age-sex group, with all of the changes being confined to the "other" component of the population.
- Beginning in January 1978, the introduction of an expansion in the sample and revisions in the estimation procedures resulted in an increase of about 250,000 in the civilian labor force and employment totals; unemployment levels and rates were essentially unchanged. An explanation of the procedural changes and an indication of the differences appear in "Revisions in the Current Population Survey in January 1978" in the February 1978 issue of this publication.
- Beginning in October 1978, the race of the individual was determined by the household respondent for the incoming rotation group households, rather than by the interviewer as before. The purpose of this change was to provide more accurate estimates of characteristics by race. Thus, in October 1978, one-eighth of the sample households had race determined by the household respondent and seven-eighths of the sample households had race determined by interviewer observation. It was not until January 1980 that the entire sample had race determined by the household respondent. The new procedure had no significant effect on the estimates.

- Beginning in January 1979, the first-stage ratio adjustment method was changed in the CPS estimation procedure. Differences between the old and new procedures existed only for metropolitan and nonmetropolitan area estimates, not for the total United States. The reasoning behind the change and an indication of the differences appear in "Revisions in the Current Population Survey in January 1979" in the February 1979 issue of this publication.
- Beginning in January 1982, the second-stage ratio adjustment method was changed. The rationale for the change and an indication of its effect on national estimates of labor force characteristics appear in "Revisions in the Current Population Survey Beginning in January 1982" in the February 1982 issue of this publication. In addition, current population estimates used in the second-stage estimation procedure were derived from information obtained from the 1980 census, rather than the 1970 census. This change caused substantial increases in the total population and in the estimates of persons in all labor force categories. Rates for labor force characteristics, however, remained virtually unchanged. Some 30,000 labor force series were adjusted back to 1970 to avoid major breaks in series. The adjustment procedure used also is described in the February 1982 article cited above. The revisions did not, however, smooth out the breaks in series occurring between 1972 and 1979 (described above), and data users should consider them when comparing estimates from different periods.
- Beginning in January 1983, the first-stage ratio adjustment method was updated to incorporate data from the 1980 census. The rationale for the change and an indication of its effect on national estimates for labor force characteristics appear in "Revisions in the Current Population Survey Beginning in January 1983" in the February 1983 issue of this publication. There were only slight differences between the old and new procedures in estimates of levels for the various labor force characteristics and virtually no differences in estimates of participation rates.
- Beginning in January 1985, most of the steps of the CPS estimation procedure—the noninterview adjustment, the first- and second-stage ratio adjustments, and the composite estimator—were revised. These procedures are described in the Estimating Methods section. A description of the changes and an indication of their effect on national estimates of labor force characteristics appear in "Changes in the Estimation Procedure in the Current Population Survey Beginning in January 1985" in the February 1985 issue of this publication. Overall, the revisions had only a slight effect on most estimates. The greatest impact was on estimates of persons of Hispanic origin. Major estimates were revised back to January 1980.
- Beginning in January 1986, the population controls used in the second-stage ratio adjustment method were revised to reflect an explicit estimate of the number of undocumented immigrants (largely Hispanic) since 1980 and an improved estimate of the number of emigrants among

- legal foreign-born residents for the same period. As a result, the total civilian population and labor force estimates were raised by nearly 400,000; civilian employment was increased by about 350,000. The Hispanic-origin population and labor force estimates were raised by about 425,000 and 305,000, respectively, and Hispanic employment was increased by 270,000. Overall and subgroup unemployment levels and rates were not significantly affected. Because of the magnitude of the adjustments for Hispanics, data were revised back to January 1980 to the extent possible. An explanation of the changes and an indication of their effect on estimates of labor force characteristics appear in "Changes in the Estimation Procedure in the Current Population Survey Beginning in January 1986" in the February 1986 issue of this publication.
- Beginning in August 1989, the second-stage ratio estimation procedures were changed slightly to decrease the chance of very small cells occurring and to be more consistent with published age, sex, race cells. This change had virtually no effect on national estimates.
- Beginning in January 1994, 1990 census-based population controls, adjusted for the estimated undercount, were introduced into the second-stage estimation procedure. This change resulted in substantial increases in total population and in all major labor force categories. Effective February 1996, these controls were introduced into the estimates for 1990-93. Under the new population controls, the civilian noninstitutional population for 1990 increased by about 1.1 million, employment by about 880,000, and unemployment by approximately 175,000. The overall unemployment rate rose by about 0.1 percentage point. For further information, see "Revisions in the Current Population Survey Effective January 1994," and "Revisions in Household Survey Data Effective February 1996" in the February 1994 and March 1996 issues, respectively, of this publication.

Additionally, for the period January through May 1994, the composite estimation procedure was suspended for technical and logistical reasons.

 Beginning in January 1997, the population controls used in the second-stage ratio adjustment method were revised to reflect updated information on the demographic characteristics of immigrants to, and emigrants from, the United States. As a result, the civilian noninstitutional population 16 years and over was raised by about 470,000. The labor force and employment levels were increased by about 320,000 and 290,000, respectively. The Hispanic-origin population and labor force estimates were raised by about 450,000 and 250,000, respectively, and Hispanic employment was increased by 325,000. Overall and subgroup unemployment rates and other percentages of labor market participation were not affected. An explanation of the changes and an indication of their effect on national labor force estimates appear in "Revisions in the Current Population Survey Effective January 1997" in the February 1997 issue of this publication.

• Beginning in January 1998, new composite estimation procedures and minor revisions in the population controls were introduced into the household survey. The new composite estimation procedures simplify processing of the monthly labor force data at BLS, allow users of the survey microdata to more easily replicate the official estimates released by BLS, and increase the reliability of the employment and labor force estimates. The new procedures also produce somewhat lower estimates of the civilian labor force and employment and slightly higher estimates of unemployment. For example, based on 1997 annual average data, the differences resulting from the use of old and new composite weights were as follows: Civilian labor force (-229,000), total employed (-256,000), and total unemployed (+27,000). Unemployment rates were not significantly affected.

Also beginning in January 1998, the population controls used in the survey were revised to reflect new estimates of legal immigration to the United States and a change in the method for projecting the emigration of foreign-born legal residents. As a result, the Hispanic-origin population was raised by about 57,000; however, the total civilian noninstitutional population 16 years and over was essentially unchanged. More detailed information on these changes and their effect on the estimates of labor force change and composition appear in "Revisions in the Current Population Survey Effective January 1998," in the February 1998 issue of this publication.

- Beginning in January 1999, the population controls used in the survey were revised to reflect newly updated information on immigration. As a result, the civilian noninstitutional population 16 years and over was raised by about 310,000. The impact of the changes varied for different demographic groups. The civilian noninstitutional population for men 16 years and over was lowered by about 185,000, while that for women was increased by about 490,000. The Hispanicorigin population was lowered by about 165,000 while that of persons of non-Hispanic origin was raised by about 470,000. Overall labor force and employment levels were increased by about 60,000 each, while the Hispanic labor force and employment estimates were reduced by about 225,000 and 215,000, respectively. The changes had only a small impact on overall and subgroup unemployment rates and other percentages of labor market participation. An explanation of the changes and an indication of their effect on national labor force estimates appear in "Revisions in the Current Population Survey Effective January 1999" in the February 1999 issue of this publication.
- Beginning in January 2003, several major changes were introduced into the CPS. These changes included:
- a) Population controls that reflected the results of Census 2000 were introduced into the monthly CPS estimation process. These new population controls substantially increased the size of the civilian noninstitutional population and the civilian labor force. Data from January 2000 through

December 2002 were revised to reflect the higher population estimates from Census 2000 and the higher rates of population growth since the census. At the start of the revision period (January 2000), the new controls raised the civilian noninstitutional population and the civilian labor force by 2.6 and 1.6 million, respectively. By December 2002, the civilian population and labor force were 3.8 and 2.5 million, respectively, higher than originally estimated. In addition to these revisions, the U.S. Census Bureau introduced another large upward adjustment to the population controls as part of its annual update of population estimates for 2003. The entire amount of this adjustment was added to the labor force data in January 2003 resulting in increases of 941,000 to the civilian noninstitutional population and 614,000 to the civilian labor force. The unemployment rate and other ratios were not substantially affected by either of these population control adjustments.

b) The modification of the questions on race and Hispanic origin to comply with new standards for maintaining, collecting, and presenting Federal data on race and ethnicity for Federal statistical agencies. In accordance with the new standards, the following changes were made to the CPS questions: 1) Individuals were now asked whether they are of Hispanic ethnicity before being asked about their race. Prior to 2003, individuals were asked their ethnic origin after they were asked about their race. 2) Individuals were now asked directly if they are Spanish, Hispanic, or Latino. Previously, individuals were identified as Hispanic based on their, or their ancestors', country of origin. 3) With respect to race, the response category of Asian and Pacific Islanders was split into two categories: a) Asian and b) Native Hawaiian or Other Pacific Islanders. 4) Individuals were allowed to choose more than one race category. Prior to 2003, individuals who considered themselves to belong to more than one race were required to select a single primary race. 5) The questions were reworded to indicate that individuals could select more than one race category and to convey more clearly that individuals should report their own perception of what their race is. These changes had no impact on the overall civilian noninstitutional population and civilian labor force but did reduce the population and labor force levels of whites, blacks or African Americans, and Asians beginning in January 2003. For whites and blacks, the differences resulted from the exclusion of individuals who reported more than one race from those groups. For Asians, the difference resulted from the same restriction as well as the split of the old Asian and Pacific Islander category into two separate categories. Analysis of data from a special CPS supplement conducted in May 2002 indicated that these changes reduced the population and labor force levels for whites by about 950,000 and 730,000, respectively, and for blacks and African Americans by about 320,000 and 240,000, respectively, while having little or no impact on their unemployment rates. For Asians, the changes had the effect

of reducing the their population by about 1.1 million and their labor force by about 720,000, but did not have a statistically significant effect on their unemployment rate. The changes did not affect the size of the Hispanic or Latino population and had no significant impact on the size of their labor force, but did cause an increase of about half a percentage point in their unemployment rate.

c) Improvements were introduced to both the secondstage and composite weighting procedures. These changes adapted the weighting procedures to the new race/ethnic classification system and enhanced the stability over time of national and State/substate labor force estimates for demographic groups.

Changes in the occupational and industrial classification systems

Beginning in 1971, the comparability of occupational employment data was affected as a result of changes in the occupational classification system for the 1970 census that were introduced into the CPS. Comparability was further affected in December 1971, when a question relating to major activity or duties was added to the monthly CPS questionnaire in order to more precisely determine the occupational classification of individuals. As a result of these changes, meaningful comparisons of occupational employment levels could not be made between 1971-72 and prior years nor between those 2 years. Unemployment rates were not significantly affected. For a further explanation of the changes in the occupational classification system, see "Revisions in Occupational Classifications for 1971" and "Revisions in the Current Population Survey" in the February 1971 and February 1972 issues, respectively, of this publication.

Beginning in January 1983, the occupational and industrial classification systems used in the 1980 census were introduced into the CPS. The 1980 census occupational classification system evolved from the Standard Occupational Classification (SOC) system and was so radically different in concepts and nomenclature from the 1970 system that comparisons of historical data are not possible without major adjustments. For example, the 1980 major group "sales occupations" is substantially larger than the 1970 category "sales workers." Major additions include "cashiers" from "clerical workers" and some self-employed proprietors in retail trade establishments from "managers and administrators, except farm."

The industrial classification system used in the 1980 census was based on the 1972 Standard Industrial Classification (SIC) system, as modified in 1977. The adoption of the new system had much less of an adverse effect on historical comparability than did the new occupational system. The most notable changes from the 1970 system were the transfer of farm equipment stores from "retail" to "wholesale" trade and of postal service from "public administration" to "transportation," and some interchange between "professional and related services" and "public administration." Additional in-

formation on the 1980 census occupational and industrial classification systems appears in "Revisions in the Current Population Survey Beginning in January 1983" in the February 1983 issue of this publication.

Beginning in January 1992, the occupational and industrial classification systems used in the 1990 census were introduced into the CPS. (These systems were based largely on the 1980 Standard Occupational Classification (SOC) and 1987 Standard Industrial Classification (SIC) systems, respectively.) There were a few breaks in comparability between the 1980 and 1990 census-based systems, particularly within the "technical, sales, and administrative support" categories. The most notable changes in industry classification were the shift of several industries from "business services" to "professional services" and the splitting of some industries into smaller, more detailed categories. A number of industry titles were changed as well, with no change in content.

Beginning in January 2003, the 2002 Census Bureau occupational and industrial classification systems were introduced into the CPS. These systems were derived from the 2000 Standard Occupational Classification (SOC) and the 2002 North American Industry Classification System (NAICS). The composition of detailed occupational and industrial classifications in the new classification systems was substantially changed from the previous systems in use as was the structure for aggregating them into broad groups. Consequently, the use of the new classification systems created breaks in existing data series at all levels of aggregation. Additional information on the 2002 Census Bureau occupational and industrial classification systems appears in "Revisions in the Current Population Survey Effective in January 2003" in the February 2003 issue of this publication.

Sampling

Since the inception of the survey, there have been various changes in the design of the CPS sample. The sample traditionally is redesigned and a new sample selected after each decennial census. Also, the number of sample areas and the number of sample persons are changed occasionally. Most of these changes are made to improve the efficiency of the sample design, increase the reliability of the sample estimates, or control cost.

Changes in this regard since 1960 are as follows: When Alaska and Hawaii received statehood in 1959 and 1960, respectively, three sample areas were added to the existing sample to account for the population of these States. In January 1978, a supplemental sample of 9,000 housing units, selected in 24 States and the District of Columbia, was designed to provide more reliable annual average estimates for States. In October 1978, a coverage improvement sample of approximately 450 sample household units representing 237,000 occupied mobile homes and 600,000 new construction housing units was added. In January 1980, another supplemental sample of 9,000 households selected in 32 States and the District of Columbia was added. A sample

reduction of about 6,000 units was implemented in May 1981. In January 1982, the sample was expanded by 100 households to provide additional coverage in counties added to the Standard Metropolitan Statistical Areas (SMSAs), which were redefined in 1973. In January 1985, a new State-based CPS sample was selected based on 1980 census information. A sample reduction of about 4,000 households was implemented in April 1988; the households were reinstated during the 8-month period, April-November 1989. A redesigned CPS sample based on the 1990 decennial census was selected for use during the 1990s. Households from this new sample were phased into the CPS between April 1994 and July 1995. The July 1995 sample was the first monthly sample based entirely on the 1990 census. For further information on the 1990 sample redesign, see "Redesign of the Sample for the Current Population Survey" in the May 1994 issue of this publication.

The original 1990 census-based sample design included about 66,000 housing units per month located in 792 selected geographic areas called primary sampling units (PSUs). The sample initially was selected to meet specific reliability criteria for the Nation, for each of the 50 States and the District of Columbia, and for the substate areas of New York City and the Los Angeles-Long Beach metropolitan area. In 1996, the original sample design reliability criteria were modified to reduce costs. In July 2001, the CPS sample was expanded to support the State Children's Health Insurance Program. For further information on the sample expansion, see "Expansion of the Current Population Survey Sample Effective July 2001" in the August 2001 issue of this publication. The current criteria, given below, are based on the coefficient of variation (CV) of the unemployment level, where the CV is defined as the standard error of the estimate divided by the estimate, expressed as a percentage. These CV controls assume a 6-percent unemployment rate to establish a consistent specification of sampling error.

The current sample design, introduced in July 2001, includes about 72,000 "assigned" housing units from 754 sample areas. Sufficient sample is allocated to maintain, at most, a 1.9-percent CV on national monthly estimates of unemployment level, assuming a 6-percent unemployment rate. This translates into a change of 0.2 percentage point in the unemployment rate being significant at a 90-percent confidence level. For each of the 50 States and for the District of Columbia, the design maintains a CV of at most 8 percent on the annual average estimate of unemployment level, assuming a 6-percent unemployment rate. About 60,000 housing units are required in order to meet the national and State reliability criteria. Due to the national reliability criterion, estimates for several large States are substantially more reliable than the State design criterion requires. Annual average unemployment estimates for California, Florida, New York, and Texas, for example, carry a CV of less than 4 percent. In support of the State Children's

Health Insurance Program, about 12,000 additional housing units are allocated to the District of Columbia and 31 States. (These are generally the States with the smallest samples after the 60,000 housing units are allocated to satisfy the national and State reliability criteria.)

In the first stage of sampling, the 754 sample areas are chosen. In the second stage, ultimate sampling unit clusters composed of about four housing units each are selected. Each month, about 72,000 housing units are assigned for data collection, of which about 60,000 are occupied and thus eligible for interview. The remainder are units found to be destroyed, vacant, converted to nonresidential use, containing persons whose usual place of residence is elsewhere, or ineligible for other reasons. Of the 60,000 housing units, about 7.5 percent are not interviewed in a given month due to temporary absence (vacation, etc.), other failures to make contact after repeated attempts, inability of persons contacted to respond, unavailability for other reasons, and refusals to cooperate (about half of the noninterviews). Information is obtained each month for about 112,000 persons 16 years of age or older.

Selection of sample areas. The entire area of the United States, consisting of 3,141 counties and independent cities, is divided into 2,007 sample units (PSUs). In most States, a PSU consists of a county or a number of contiguous counties. In New England and Hawaii, minor civil divisions are used instead of counties.

Metropolitan areas within a State are used as a basis for forming PSUs. Outside of metropolitan areas, counties normally are combined except when the geographic area of an individual county is too large. Combining counties to form PSUs provides greater heterogeneity; a typical PSU includes urban and rural residents of both high and low economic levels and encompasses, to the extent feasible, diverse occupations and industries. Another important consideration is that the PSU be sufficiently compact so that, with a small sample spread throughout, it can be efficiently canvassed without undue travel cost.

The 2,007 PSUs are grouped into strata within each State. Then, one PSU is selected from each stratum with the probability of selection proportional to the population of the PSU. Nationally, there are a total of 428 PSUs in strata by themselves. These strata are self-representing and are generally the most populous PSUs in each State. The 326 remaining strata are formed by combining PSUs that are similar in such characteristics as unemployment, proportion of housing units with three or more persons, number of persons employed in various industries, and average monthly wages for various industries. The single PSU randomly selected from each of these strata is nonself-representing because it represents not only itself but the entire stratum. The probability of selecting a particular PSU in a nonself-representing stratum is proportional to its 1990 population. For

example, within a stratum, the chance that a PSU with a population of 50,000 would be selected for the sample is twice that for a PSU having a population of 25,000.

Selection of sample households. Because the sample design is State based, the sampling ratio differs by State and depends on State population size as well as both national and State reliability requirements. The State sampling ratios range roughly from 1 in every 100 households to 1 in every 3,000 households. The sampling ratio occasionally is modified slightly to hold the size of the sample relatively constant given the overall growth of the population. The sampling ratio used within a sample PSU depends on the probability of selection of the PSU and the sampling ratio for the State. In a sample PSU with a probability of selection of 1 in 10 and a State sampling ratio of 3,000, a within-PSU sampling ratio of 1 in 300 achieves the desired ratio of 1 in 3,000 for the stratum.

The 1990 within-PSU sample design was developed using block-level data from the 1990 census. (The 1990 census was the first decennial census that produced data at the block level for the entire country.) Normally, census blocks are bounded by streets and other prominent physical features such as rivers or railroad tracks. County, minor civil division, and census place limits also serve as block boundaries. In cities, blocks can be bounded by four streets and be quite small in land area. In rural areas, blocks can be several square miles in size.

For the purpose of sample selection, census blocks were grouped into three strata: Unit, group quarters, and area. (Occasionally, units within a block were split between the unit and group-quarters strata.) The unit stratum contained regular housing units with addresses that were easy to locate (for example, most single-family homes, townhouses, condominiums, apartment units, and mobile homes). The group-quarters stratum contained housing units in which residents shared common facilities or received formal or authorized care or custody. Unit and group-quarters blocks exist primarily in urban areas. The area stratum contains blocks with addresses that are more difficult to locate. Area blocks exist primarily in rural areas.

To reduce the variability of the survey estimates and to ensure that the within-PSU sample would reflect the demographic and socioeconomic characteristics of the PSU, blocks within the unit, group-quarters, and area strata were sorted using geographic and block-level data from the census. Examples of the census variables used for sorting include proportion of minority renter-occupied housing units, proportion of housing units with female householders, and proportion of owner-occupied housing units. The specific sorting variables used differed by type of PSU (urban or rural) and stratum.

Within each block, housing units were sorted geographically and grouped into clusters of approximately four units. A systematic sample of these clusters was then selected independently from each stratum using the appropriate within-PSU sampling ratio. The geographic clustering of the sample units reduces field representative travel costs. Prior to inter-

viewing, special listing procedures are used to locate the particular sample addresses in the group-quarters and area blocks.

Units in the three strata described above all existed at the time of the 1990 decennial census. Through a series of additional procedures, a sample of building permits is included in the CPS to represent housing units built after the decennial census. Adding these newly built units keeps the sample up-to-date and representative of the population. It also helps to keep the sample size stable: Over the life of the sample, the addition of newly built housing units compensates for the loss of "old" units that may be abandoned, demolished, or converted to nonresidential use.

Rotation of sample. Part of the sample is changed each month. Each monthly sample is divided into eight representative subsamples or rotation groups. A given rotation group is interviewed for a total of 8 months, divided into two equal periods. It is in the sample for 4 consecutive months, leaves the sample during the following 8 months, and then returns for another 4 consecutive months. In each monthly sample, one of the eight rotation groups is in the first month of enumeration, another rotation group is in the second month, and so on. Under this system, 75 percent of the sample is common from month to month, and 50 percent is common from year to year for the same month. This procedure provides a substantial amount of month-to-month and year-to-year overlap in the sample, thus providing better estimates of change and reducing discontinuities in the data series without burdening any specific group of households with an unduly long period of inquiry.

CPS sample, 1947 to present. Table 1-A provides a description of some aspects of the CPS sample designs in use since 1947. A more detailed account of the history of the CPS sample design appears in chapter 2 of "The Current Population Survey: Design and Methodology," Technical Paper 63RV, (Washington, U.S. Census Bureau and Bureau of Labor Statistics, March 2002), available on the Internet at www.bls.census.gov/cps/tp/tp63.htm. A description of the 1990 census-based sample design appears in "Redesign of the Sample for the Current Population Survey," in the May 1994 issue of this publication and in chapter 3 of Technical Paper 63RV referenced above. A description of the sample expansion in support of the State Children's Health Insurance Program appears in "Expansion of the Current Population Survey Sample Effective July 2001", in the August 2001 issue of this publication and in Appendix J, "Changes to the Current Population Survey Sample in July 2001," of Technical Paper 63RV referenced above.

ESTIMATING METHODS

Under the estimating methods used in the CPS, all of the results for a given month become available simultaneously and are based on returns from the entire panel of respondents. The estimation procedure involves weighting the data from each sample person by the inverse of the probability of the person being in the sample. This gives a rough measure

Table 1-A. Characteristics of the CPS sample, 1947 to present

Period	Number of sample	Househol	Households visited		
	areas	Interviewed	Not interviewed	but not eligible	
Aug. 1947 to Jan. 1954 Feb. 1954 to Apr. 1956 May 1956 to Dec. 1959 Jan. 1960 to Feb. 1963 Mar. 1963 to Dec. 1966 Jan. 1967 to July 1971 Aug. 1971 to July 1972 Aug. 1972 to Dec. 1977 Jan. 1978 to Dec. 1979 Jan. 1980 to Apr. 1981 May 1981 to Dec. 1984 Jan. 1985 to Mar. 1988 Apr. 1988 to Mar. 1988 Apr. 1989 to Oct. 1994 3 Nov. 1994 to Aug. 1995 4 Sept. 1995 to Dec. 1995 Jan. 1996 to June 2001	68 230 1330 2333 357 449 449 461 614 629 629 729 729 729 729 792 792	21,000 21,000 33,500 33,500 48,000 45,000 45,000 53,500 62,200 57,800 57,000 53,200 57,400 54,500 52,900 46,250	500-1,000 500-1,000 1,500 1,500 2,000 2,000 2,000 2,500 2,800 2,500 2,500 2,600 2,600 3,500 3,400 3,750	3,000-3,500 3,000-3,500 6,000 6,000 8,500 8,000 10,000 12,000 11,000 11,500 11,800 10,000 9,700 10,000	
July 2001 to present ⁵	754	55,500	4,500	12,000	

Beginning in May 1956, these areas were chosen to provide coverage in each State and the District of Columbia.

of the number of actual persons that the sample person represents. Since 1985, most sample persons within the same State have had the same probability of selection. Some selection probabilities may differ within a State due to the sample design or for operational reasons. Field subsampling, for example, which is carried out when areas selected for the sample are found to contain many more households than expected, may cause probabilities of selection to differ for some sample areas within a State. Through a series of estimation steps (outlined below), the selection probabilities are adjusted for noninterviews and survey undercoverage; data from previous months are incorporated into the estimates through the composite estimation procedure.

1. Noninterview adjustment. The weights for all interviewed households are adjusted to account for occupied sample households for which no information was obtained because of absence, impassable roads, refusals, or unavailability of the respondents for other reasons. This noninterview adjustment is made separately for clusters of similar sample areas that are usually, but not necessarily, contained within a State. Similarity of sample areas is based on Metropolitan Statistical Area (MSA) status and size. Within each cluster, there is a further breakdown by residence. Each MSA cluster is split by "central city" and "balance of the MSA." Each non-MSA cluster is split by "urban" and "rural" residence categories. The proportion of sample households not interviewed varies from 7 to 8 percent, depending on weather, vacation, etc.

2. Ratio estimates. The distribution of the population selected for the sample may differ somewhat, by chance, from that of the population as a whole in such characteristics as age, race, sex, and State of residence. Because these charac-

teristics are closely correlated with labor force participation and other principal measurements made from the sample, the survey estimates can be substantially improved when weighted appropriately by the known distribution of these population characteristics. This is accomplished through two stages of ratio adjustment, as follows:

- a. First-stage ratio estimation. The purpose of the first-stage ratio adjustment is to reduce the contribution to variance that results from selecting a sample of PSUs rather than drawing sample households from every PSU in the Nation. This adjustment is made to the CPS weights in two race cells: Black and nonblack; it is applied only to PSUs that are not self-representing and for those States that have a substantial number of black households. The procedure corrects for differences that existed in each State cell at the time of the 1990 census between 1) the race distribution of the population in sample PSUs and 2) the race distribution of all PSUs. (Both 1 and 2 exclude self-representing PSUs.)
- **b.** Second-stage ratio estimation. This procedure substantially reduces the variability of estimates and corrects, to some extent, for CPS undercoverage. A national-coverage step and a State-coverage step make preliminary corrections for undercoverage. The CPS sample weights are then adjusted to ensure that sample-based estimates of population match independent population controls. Three sets of controls are used in different steps of the procedure:
- *1) State step:* Civilian noninstitutional population controls for 6 age-sex cells in the Los Angeles-Long Beach metropolitan area, the balance of California, New York City, the balance of New York State, each of the other 48 States, and the District of Columbia.

 $^{^{\,2}\,}$ Three sample areas were added in 1960 to represent Alaska and Hawaii after statehood.

³ The sample was increased incrementally during the 8-month period, April-November 1989.

⁴ Includes 2,000 additional assigned housing units from Georgia and Virginia that were gradually phased in during the 10-month period, October 1994-August 1995.

⁵ Includes 12,000 assigned housing units in support of the State Children's Health Insurance Program.

- *2) Ethnicity step:* National civilian noninstitutional population controls for 26 Hispanic and 26 non-Hispanic age-sex cells.
- *3) Race step:* National civilian noninstitutional population controls for 34 white, 26 black, and 26 Asian-plus-residual-race age-sex cells.

The independent population controls are prepared by projecting forward the resident population as enumerated on April 1, 2000. The projections are derived by updating demographic census data with information from a variety of other data sources that account for births, deaths, and net migration. Estimated numbers of resident Armed Forces personnel and institutionalized persons reduce the resident population to the civilian noninstitutional population. Prior to January 2003, the projections were based on earlier censuses. See "Revisions to the Current Population Survey Effective in January 2003," in the February 2003 issue of this publication for a detailed discussion of changes to the second-stage weighting and composite estimating procedures that were introduced in January 2003.

3. Composite estimation procedure. The last step in the preparation of most CPS estimates makes use of a composite estimation procedure. The composite estimate consists of a weighted average of two factors: The two-stage ratio estimate based on the entire sample from the current month and the composite estimate for the previous month, plus an estimate of the month-to-month change based on the six rotation groups common to both months. In addition, a bias adjustment term is added to the weighted average to account for relative bias associated with month-in-sample estimates. This month-in-sample bias is exhibited by unemployment estimates for persons in their first and fifth months in the CPS being generally higher than estimates obtained for the other months.

The composite estimate results in a reduction in the sampling error beyond that which is achieved after the two stages of ratio adjustment. For some items, the reduction is substantial. The resultant gains in reliability are greatest in estimates of month-to-month change, although gains usually are also obtained for estimates of level in a given month, change from year to year, and change over other intervals of time.

Rounding of estimates

The sums of individual items may not always equal the totals shown in the same tables because of independent rounding of totals and components to the nearest thousand. Similarly, sums of percent distributions may not always equal 100 percent because of rounding. Differences, however, are insignificant.

Reliability of the estimates

An estimate based on a sample survey has two types of error sampling error and nonsampling error. The estimated standard errors provided in this publication are approximations

of the true sampling errors. They incorporate the effect of some nonsampling errors in response and enumeration, but do not account for any systematic biases in the data.

Nonsampling error. The full extent of nonsampling error is unknown, but special studies have been conducted to quantify some sources of nonsampling error in the CPS. The effect of nonsampling error is small on estimates of relative change, such as month-to-month change; estimates of monthly levels tend to be affected to a greater degree.

Nonsampling errors in surveys can be attributed to many sources, for example, the inability to obtain information about all persons in the sample; differences in the interpretation of questions; inability or unwillingness of respondents to provide correct information; inability of respondents to recall information; errors made in collecting and processing the data; errors made in estimating values for missing data; and failure to represent all sample households and all persons within sample households (undercoverage).

Nonsampling errors occurring in the interview phase of the survey are studied by means of a reinterview program. This program is used to estimate various sources of error, as well as to evaluate and control the work of the interviewers. A random sample of each interviewer's work is inspected through reinterview at regular intervals. The results indicate, among other things, that the data published from the CPS are subject to moderate systematic biases. A description of the CPS reinterview program may be found in Appendix G, "Reinterview: Design and Methodology," of "The Current Population Survey: Design and Methodology," Technical Paper 63RV (Washington, U.S. Census Bureau and Bureau of Labor Statistics, March 2002), available on the Internet at www.bls.census.gov/cps/tp/tp63.htm.

The effects of some components of nonsampling error in the CPS data can be examined as a result of the rotation plan used for the sample, because the level of the estimates varies by rotation group. A description appears in Barbara A. Bailar, "The Effects of Rotation Group Bias on Estimates from Panel Surveys," *Journal of the American Statistical Association*, March 1975, pp. 23-30.

Undercoverage in the CPS results from missed housing units and missed persons within sample households. The CPS covers about 92 percent of the decennial census population (adjusted for census undercount). It is known that the CPS undercoverage varies with age, sex, race, and Hispanic origin. Generally, undercoverage is larger for men than for women and is larger for blacks, Hispanics, and other races than for whites. Ratio adjustment to independent age-sex-race-origin population controls, as described previously, partially corrects for the biases due to survey undercoverage. However, biases exist in the estimates to the extent that missed persons in missed households or missed persons in interviewed households have characteristics different from those of interviewed persons in the same age-sex-race-origin group.

Additional information on nonsampling error in the CPS

appears in Camilla Brooks and Barbara Bailar, "An Error Profile: Employment as Measured by the Current Population Survey," Statistical Policy Working Paper 3 (Washington, U.S. Department of Commerce, Office of Federal Statistical Policy and Standards, September 1978); Marvin Thompson and Gary Shapiro, "The Current Population Survey: An Overview," *Annals of Economic and Social Measurement*, Vol. 2, April 1973; and "The Current Population Survey: Design and Methodology," Technical Paper 63RV referenced above. The last document includes a comprehensive discussion of various sources of errors and describes attempts to measure them in the CPS.

Sampling error. When a sample, rather than the entire population, is surveyed, estimates differ from the true population values that they represent. This difference, or sampling error, occurs by chance, and its variability is measured by the standard error of the estimate. Sample estimates from a given survey design are unbiased when an average of the estimates from all possible samples would yield, hypothetically, the true population value. In this case, the sample estimate and its standard error can be used to construct approximate confidence intervals, or ranges of values that include the true population value with known probabilities. If the process of selecting a sample from the population were repeated many times, an estimate made from each sample, and a suitable estimate of its standard error calculated for each sample, then:

- 1. Approximately 68 percent of the intervals from one standard error below the estimate to one standard error above the estimate would include the true population value.
- 2. Approximately 90 percent of the intervals from 1.645 standard errors below the estimate to 1.645 standard errors above the estimate would include the true population value.
- 3. Approximately 95 percent of the intervals from 1.96 standard errors below the estimate to 1.96 standard errors above the estimate would include the true population value.

These confidence interval statements are approximately true for the CPS. Although the estimating methods used in the CPS do not produce unbiased estimates, biases for most estimates are believed to be small. Methods for estimating standard errors reflect not only sampling errors but also some kinds of nonsampling error. Although both the estimates and the estimated standard errors depart from the theoretical ideal, the departures are minor and have little impact on the confidence interval statements. When clarity is needed, an estimated confidence interval is specified to be "approximate," as is the estimated standard error used in the computation.

Tables 1-B through 1-D are provided so that approximate standard errors of estimates can be easily obtained. Tables 1-B and 1-C give approximate standard errors for estimated monthly levels and rates for selected employment status characteristics; the tables also provide approximate standard errors for consecutive month-to-month changes in the estimates. It is impractical to show approximate standard errors for all CPS estimates in this publication, so table 1-D provides parameters and factors that allow the user to calculate approximate standard errors for a wide range of estimated levels, rates, and percentages, and also changes over time. The parameters and factors are used in formulas that are commonly called *generalized variance functions*.

The approximate standard errors provided in this publication are based on the sample design and estimation procedures as of 1996, and reflect the population levels and sample size as of that year. Standard errors for years prior to 1996 may be roughly approximated by applying these adjustments

 $\label{thm:continuous} \mbox{Table 1-B. } \mbox{\bf Approximate standard errors for major employment status categories}$

(In thousands)

,		
Characteristic	Monthly level	Consecutive month-to- month change
Total		
Total, 16 years and over: Civilian labor force Employed Unemployed	267 273 131	174 177 166
Men, 20 years and over: Civilian labor force Employed Unemployed	184 196 83	120 128 106
Women, 20 years and over: Civilian labor force Employed Unemployed	209 215 77	136 140 98
Both sexes, 16 to 19 years: Civilian labor force Employed Unemployed	90 95 56	87 91 93
Black or African American		
Total, 16 years and over: Civilian labor force Employed Unemployed	113 121 64	73 79 81
Men, 20 years and over: Civilian labor force Employed Unemployed	81 85 39	53 55 50
Women, 20 years and over: Civilian labor force Employed Unemployed	72 77 40	47 50 50
Both sexes, 16 to 19 years: Civilian labor force Employed Unemployed	42 39 28	40 38 46
Hispanic or Latino ethnicity		
Total, 16 years and over: Civilian labor force Employed Unemployed	90 100 54	59 65 69

Table 1-C. Approximate standard errors for unemployment rates by major characteristics

(In percent)

Characteristic	Monthly rate	Consecutive month-to- month change
Total	0.09 .12 .12 .13 .13	0.12 .16 .15 .17 .16 1.08
White	.10 .39 .37	.12 .49 .47
Married men, spouse present	.12 .14 .43	.15 .18 .54

to the standard errors presented here. (More accurate standard error estimates for historical CPS data may be found in previous issues of this publication.)

- 1. For the years 1967 through 1995, multiply the standard errors by 0.96.
- 2. For the years 1956 through 1966, multiply the standard errors by 1.17.
- 3. For years prior to 1956, multiply the standard errors by 1.44.

Use of tables 1-B and 1-C. These tables provide a quick reference for standard errors of major characteristics. Table 1-B gives approximate standard errors for estimates of monthly levels and consecutive month-to-month changes in levels for major employment status categories. Table 1-C gives approximate standard errors for estimates of monthly unemployment rates and consecutive month-to-month changes in unemployment rates for some demographic, occupational, and industrial categories. For characteristics not given in tables 1-B and 1-C, refer to table 1-D.

Illustration. Suppose that, for a given month, the number of women age 20 years and over in the civilian labor force is estimated to be 60,000,000. For this characteristic, the approximate standard error of 209,000 is given in table 1-B in the row "Women, 20 years and over; Civilian labor force." To calculate an approximate 90-percent confidence interval, multiply the standard error of 209,000 by the factor 1.645 to obtain 344,000. This number is subtracted from and then added to 60,000,000 to obtain an approximate 90-percent confidence interval: 59,656,000 to 60,344,000. Concluding that the true civilian labor force level lies within an interval calculated in this way would be correct for roughly 90 percent of all possible samples that could have been selected for the CPS.

Use of table 1-D. This table gives **a** and **b** parameters that can be used with formulas to calculate approximate monthly

standard errors for a wide range of estimated levels, proportions, and rates. Factors are provided to convert monthly measures into approximate standard errors of estimates for other periods (quarterly and yearly averages) and approximate standard errors for changes over time (consecutive monthly changes, changes in consecutive quarterly and yearly averages, and changes in monthly estimates 1 year apart).

The standard errors for estimated changes in level from one month to the next, one year to the next, etc., depend more on the monthly levels for characteristics than on the size of the changes. Likewise, the standard errors for changes in rates (or percentages) depend more on the monthly rates (or percentages) than on the size of the changes. Accordingly, the factors presented in table 1-D are applied to the monthly standard error approximations for levels, percentages, or rates; the magnitudes of the changes do not come into play. Factors are not given for estimated changes between nonconsecutive months (except for changes of monthly estimates 1 year apart); however, the standard errors may be assumed to be higher than the standard errors for consecutive monthly changes.

$$se(x) = \sqrt{ax^2 + bx}$$

Standard errors of estimated levels using table 1-D. The approximate standard error se(x) of x, an estimated monthly level, can be obtained using the formula below, where a and b are the parameters from table 1-D associated with a particular characteristic.

Illustration. Assume that, in a given a month, there are an estimated 3 million unemployed men. Obtain the appropriate \boldsymbol{a} and \boldsymbol{b} parameters from table 1-D (Total or white; Men; Unemployed). Use the formula for $\boldsymbol{se}(\boldsymbol{x})$ to compute an approximate standard error on the estimate of $\boldsymbol{x} = 3,000,000$.

$$\mathbf{a} = -0.0000348$$
 $\mathbf{b} = 2927.43$

$$se(3,000,000) = \sqrt{-0.0000348(3,000,000)^2 + 2927.43(3,000,000)} \approx 92,000$$

Procedure for using table 1-D factors for levels. Table 1-D gives factors that can be used to compute approximate standard errors of levels for other periods or for changes over time. For each characteristic, factors **f** are given for:

Consecutive month-to-month changes

Changes in monthly estimates 1 year apart

Quarterly averages

Changes in consecutive quarterly averages

Yearly averages

Changes in consecutive yearly averages

For a given characteristic, the table 1-D factor is used in the following formula, which also uses the **a** and **b** parameters from the same line of the table. A three-step procedure for using the formula is given. The **f** in the formula is frequently called an **adjustment factor**; because it appears to adjust a monthly standard error **se(x)**. However, the **x** in the formula is not a monthly level, but an average of several monthly levels (see examples listed under Step 1, below).

$$se(x, f) = f * se(x) = f * \sqrt{(ax^2 + bx)}$$

where x is an average of monthly levels over a designated period.

Step 1. Average monthly levels appropriately in order to obtain **x**. Levels for 3 months are averaged for quarterly averages, and those for 12 months are averaged for yearly averages. For changes in consecutive averages, average over the 2 months, 2 quarters, or 2 years involved. For changes in monthly estimates 1 year apart, average the 2 months involved.

Step 2. Calculate an approximate standard error se(x), treating the average x from step 1 as if it were an estimate of level for a single month. Obtain parameters a and b from table 1-D. (Note that, for some characteristics, an approximate standard error of level could instead be obtained from table 1-B and used in place of se(x) in the formula.)

Step 3. Determine the standard error se(x, f) on the average level or on the change in level. Multiply the result from step 2 by the appropriate factor f. The f and f parameters used in step 2 and the factor f used in this step come from the same line in table 1-D.

Illustration of a standard error computation for consecutive month change in level. Continuing the previous example, suppose that in the next month the estimated number of unemployed men increases by 150,000, from 3,000,000 to 3,150,000.

Step 1. The average of the two monthly levels is x = 3,075,000.

Step 2. Apply the **a** and **b** parameters from table 1-D (Total or white; Men; Unemployed) to the average **x**, treating it like an estimate for a single month.

$$a = -0.0000348$$
 $b = 2927.43$

$$se(3,075,000) = \sqrt{-0.0000348(3,075,000)^2 + 2927.43(3,075,000)} \approx 93,000$$

Step 3. Obtain f = 1.27 from the same row of table 1-D in the column "Consecutive month-to-month change," and multiply the factor by the result from step 2.

$$se(150,000) = f * se(3,075,000) = 1.27 * 93,000 \approx 118,000$$

For an approximate 90-percent confidence interval, compute $1.645 * 118,000 \approx 194,000$. Subtract the number from and add the number to 150,000 to obtain an interval of -44,000 to 344,000. This is an approximate 90-percent confidence interval for the true change, and since this interval includes zero, one cannot assert at this level of confidence that any real change has occurred in the unemployment level. The result also can be expressed by saying that the apparent change of 150,000 is not significant at a 90-percent confidence level.

Illustration of a standard error computation for quarterly average level. Suppose that an approximate standard error is desired for a quarterly average of the black or Africian American employment level. Suppose that the estimated employment levels for the 3 months making up the quarter are 14,900,000, 15,000,000, and 15,100,000.

Step 1. The average of the three monthly levels is x = 15,000,000.

Step 2. Apply the \boldsymbol{a} and \boldsymbol{b} parameters from table 1-D (Black; Total; Civilian labor force, employed, and not in labor force) to the average \boldsymbol{x} , treating it like an estimate for a single month.

$$\mathbf{a} = -0.0001541$$
 $\mathbf{b} = 3295.99$

$$se(15,000,000) = \sqrt{-0.0001541(15,000,000)^2 + 3295.99(15,000,000)} \approx 122,000$$

Step 3. Obtain f= .86 from the same row of table 1-D in the column "Quarterly averages," and multiply the factor by the result from step 2.

$$se(15,000,000) = .86 *122,000 \approx 105,000$$

Illustration of a standard error computation for change in quarterly level. Continuing the example, suppose that, in the next quarter, the estimated average employment level for blacks is 15,400,000, based on monthly levels of 15,300,000, 15,400,000, and 15,500,000. This is an estimated increase of 400,000 over the previous quarter.

Step 1. The average of the two quarterly levels is x = 15,200,000.

Step 2. Apply the **a** and **b** parameters from table 1-D (Black; Total; Civilian labor force, employed, and not in labor force) to the average **x**, treating it like an estimate for a single month.

$$\mathbf{a} = -0.0001541$$
 $\mathbf{b} = 3295.99$

$$se(15,200,000) = \sqrt{-0.0001541(15,200,000)^2 + 3295.99(15,200,000)} \approx 120,000$$

Step 3. Obtain \mathbf{f} = .78 from the same row of table 1-D in the column "Change in consecutive quarterly averages," and multiply the factor by the result from step 2.

$$se(400,000) = .78 * se(15,200,000) = .78 * 120,000 \approx 94,000$$

For an approximate 95-percent confidence interval, compute $1.96 * 94,000 \approx 184,000$. Subtract the number from and add the number to 400,000 to obtain an interval of 216,000 to 584,000. The interval excludes zero. Another way of stating this is to observe that the estimated change of 400,000 clearly exceeds 1.96 standard errors, or 184,000. One can conclude from these data that the change in quarterly averages is significant at a 95-percent confidence level.

Standard errors of estimated rates and percentages using table 1-D. As shown in the formula below, the approximate standard error se(p,y) of an estimated rate or percentage p depends, in part, upon the number of persons y in its base or denominator. Generally, rates and percentages are not published unless the monthly base is greater than 75,000 persons, the quarterly average base is greater than 60,000 persons, or the yearly average base is greater than 35,000 persons. The b parameter is obtained from table 1-D. When the base y and the numerator of p are from different categories within the table, use the p parameter from table 1-D relevant to the numerator of the rate or percentage.

$$se(p, y) = \sqrt{\frac{b}{y}p(100 - p)}$$

Note that se(p, y) is in percent.

Illustration. For a given month, suppose y = 6,200,000 women 20 to 24 years of age are estimated to be employed. Of this total, 2,000,000, or p = 32 percent, are classified as part-time workers. Obtain the parameter b = 3005.06 from the table 1-D row (Employment; Part-time workers) that is relevant to the numerator of the percentage. Apply the formula to obtain:

$$se(p, y) = \sqrt{\frac{3005.06}{6,200,000}(32)(100 - 32)} \approx 1.0$$
 percent

For an approximate 95-percent confidence interval, compute 1.96 * 1.0 percent, and round the result to 2 percent. Subtract this from and add this to the estimate of p = 32 percent to obtain an interval of 30 percent to 34 percent.

Procedure for using table 1-D factors for rates and percentages. Table 1-D factors can be used to compute approximate standard errors on rates and percentages for other periods or for changes over time. As for levels, there are three steps in the procedure for using the formula.

$$se(p, y, f) = f * se(p, y) = f * \sqrt{\frac{b}{y} p(100 - p)}$$

where p and y are averages of monthly estimates over a designated period. Note that se(p, y, f) is in percent.

Step 1. Appropriately average estimates of monthly rates or percentages to obtain **p**, and also average estimates of monthly levels to obtain **y**. Rates for 3 months are averaged for quarterly averages, and those for 12 months are averaged for yearly averages. For changes in consecutive averages, average over the 2 months, 2 quarters, or 2 years involved. For changes in monthly estimates 1 year apart, average the 2 months involved.

Step 2. Calculate an approximate standard error se(p, y), treating the averages p and y from step 1 as if they were estimates for a single month. Obtain the b parameter from the table 1-D row that describes the numerator of the rate or percentage. (Note that, for some characteristics, an approximate standard error could instead be obtained from table 1-C and used in place of se(p, y) in the formula.)

Step 3. Determine the standard error se(p, y, f) on the average level or on the change in level. Multiply the result from step 2 by the appropriate factor f. The b parameter used in step 2 and the factor f used in this step come from the same line in table 1-D.

Illustration of a standard error computation for consecutive month change in percentage. Continuing the previous example, suppose that, in the next month, 6,300,000 women 20 to 24 years of age are reported employed, and that 2,150,000, or 34 percent, are part-time workers.

Step 1. The month-to-month change is 2 percent = 34 percent - 32 percent. The average of the two monthly percentages of 32 percent and 34 percent is needed (p = 33 percent), as is the average of the two bases of 6,200,000 and 6,300,000 (y = 6,250,000).

Step 2. Apply the b = 3005.06 parameter from table 1-D (Employment; Part-time workers) to the averaged p and y, treating the averages like estimates for a single month.

$$se(p, y) = \sqrt{\frac{3005.06}{6.250.000}(33)(100 - 33)} \approx 1.0$$
 percent

Step 3. Obtain f=.65 from the same row of table 1-D in the column "Consecutive month-to-month change," and multiply the factor by the result from step 2.

$$se(2\%) = .65*1.0 \text{ percent} = .65 \text{ percent}$$

For an approximate 95-percent confidence interval, compute 1.96 * .65 percent, and round the result to 1.3 percent. Subtract this from and add this to the 2-percent estimate of change to obtain an interval of 0.7 percent to 3.3 percent. Because this interval excludes zero, it can be concluded at a 95-percent confidence level that the change is significant.

 $\label{thm:computation} \textbf{Table 1-D. Parameters and factors for computation of approximate standard errors for estimates of monthly levels}$

	Parameters		Factors					
Characteristic	а	b	Consecutive month-to- month change	Year-to-year change of monthly estimates	Quarterly averages	Change in consecutive quarterly averages	Yearly averages	Change in consecutive yearly averages
Total or white								
Total: Civilian labor force, employed, and not in labor force Unemployed	-0.0000077	1586.29	0.65	1.22	0.87	0.77	0.68	0.81
	0000174	3005.06	1.27	1.38	.72	.91	.42	.57
Men: Civilian labor force, employed, and not in labor force Unemployed	0000348	2927.43	.65	1.23	.86	.79	.66	.80
	0000348	2927.43	1.27	1.39	.72	.91	.43	.57
Women: Civilian labor force, employed, and not in labor force Unemployed	0000325	2693.27	.65	1.22	.87	.78	.67	.81
	0000325	2693.27	1.27	1.39	.71	.90	.41	.55
Both sexes, 16 to 19 years: Civilian labor force, employed, and not in labor force Unemployed	0002436 0002436	3005.06 3005.06	.96 1.65	1.32 1.37	.81 .68	.87 .88	.55 .40	.71 .53
Black or Africian American								
Total: Civilian labor force, employed, and not in labor force Unemployed	0001541	3295.99	.65	1.22	.86	.78	.66	.80
	0001541	3295.99	1.28	1.38	.73	.90	.43	.58
Men: Civilian labor force, employed, and not in labor force Unemployed	0003361	3332.28	.65	1.25	.84	.82	.62	.76
	0003361	3332.28	1.27	1.37	.73	.91	.43	.58
Women: Civilian labor force, employed, and not in labor force Unemployed	0002821	2944.26	.65	1.27	.84	.80	.64	.78
	0002821	2944.26	1.27	1.39	.71	.90	.41	.56
Both sexes, 16 to 19 years: Civilian labor force, employed, and not in labor force Unemployed	0015306	3295.99	.96	1.33	.80	.85	.56	.70
	0015306	3295.99	1.65	1.37	.68	.86	.41	.52
Hispanic or Latino ethnicity								
Total: Civilian labor force, employed, and not in labor force Unemployed	0001260	3295.99	.65	1.20	.86	.82	.65	.78
	0001260	3295.99	1.28	1.38	.71	.90	.42	.56
Men: Civilian labor force, employed, and not in labor force Unemployed	0002570	3332.28	.65	1.26	.84	.82	.62	.76
	0002570	3332.28	1.29	1.38	.71	.90	.41	.55
Women: Civilian labor force, employed, and not in labor force Unemployed	0002140	2944.26	.65	1.21	.86	.84	.63	.76
	0002140	2944.26	1.27	1.38	.71	.89	.41	.55
Both sexes, 16 to 19 years: Civilian labor force, employed, and not in labor force Unemployed	0014250 0014250	3295.99 3295.99	.96 1.65	1.34 1.42	.81 .70	.84 .89	.58 .41	.73 .55

Table 1-D. Parameters and factors for computation of approximate standard errors for estimates of monthly levels—Continued

	Parameters		Factors					
Characteristic	а	b	Consecutive month-to- month change	Year-to-year change of monthly estimates	Quarterly averages	Change in consecutive quarterly averages	Yearly averages	Change in consecutive yearly averages
Employment								
Educational attainment	-0.0000174	3005.06	0.65	1.11	0.87	0.92	0.61	0.74
Marital status, men Marital status, women Women who maintain families .	.0000348.0000325.0000325	2927.43 2693.27 2693.27	.65 .65 .65	1.15 1.18 1.18	.86 .85 .85	.93 .94 .94	.59 .57 .57	.72 .72 .72
Nonagricultural industries: Total Wage and salary workers Self-employed workers Unpaid family workers	0000174 0000174 0000174 0000174	3005.06 3005.06 3005.06 3005.06	.65 .65 .65 .65	1.15 1.13 1.15 1.26	.88 .88 .87 .81	.75 .84 .96 .95	.71 .67 .58 .50	.83 .79 .71 .65
Full-time workers	0000174 0000174	3005.06 3005.06	.65 .65	1.17 1.27	.85 .81	.92 .89	.59 .55	.72 .69
Multiple jobholders	0000174	3005.06	1.27	1.29	.78	.91	.50	.64
At work								
Total and nonagricultural industries: Total	0000174 0000174 0000174 0000174 0000174 0000174 0000174 0000174	3005.06 3005.06 3005.06 3005.06 3005.06 3005.06 3005.06 3005.06	.65 1.65 1.27 1.65 1.27 1.65 1.27 1.47	1.21 1.36 1.33 1.34 1.30 1.34 1.25 1.37	.84 .67 .73 .67 .76 .71 .78 .67	.77 .86 .88 .86 .87 .86 .86	.66 .38 .45 .39 .51 .45 .53 .39	.79 .51 .58 .51 .64 .57 .65 .52
Unemployment								
Educational attainment	0000174	3005.06	1.27	1.38	.72	.91	.42	.57
Marital status, men Marital status, women Women who maintain families	0000348 0000325 0000325	2927.43 2693.27 2693.27	1.27 1.27 1.27	1.39 1.39 1.39	.72 .71 .71	.91 .90 .90	.43 .41 .41	.57 .55 .55
Industries and occupations	0000174	3005.06	1.27	1.38	.72	.91	.42	.57
Full-time workers	0000174 0000174	3005.06 3005.06	1.27 1.65	1.38 1.40	.72 .69	.91 .88	.42 .40	.57 .53
Less than 5 weeks	0000174 0000174 0000174 0000174	3005.06 3005.06 3005.06 3005.06	1.27 1.65 1.65 1.27	1.38 1.37 1.39 1.42	.72 .66 .67 .75	.91 .88 .89 .93	.42 .35 .36 .44	.57 .50 .50 .60
All reasons for unemployment, except temporary layoff On temporary layoff	0000174 0000174	3005.06 3005.06	1.27 1.65	1.38 1.35	.72 .68	.91 .87	.42 .40	.57 .53
Not in the labor force								
Total Persons who currently want a job and discouraged	0000077	1586.29	.65	1.22	.87	.77	.68	.81
workers	0000174	3005.06	1.65	1.41	.63	.83	.36	.48